

PATENT

ATTY DKT. P02039US1

APPLICATION FOR UNITED STATES LETTERS PATENT

FOR

DNA TEST FOR SCA10

By

Tetsuo Ashizawa

And

Tohru Matsuura

TOHRO MATSUURA

DNA TEST FOR SCA10

[0001] The U.S. Government has rights in the invention by virtue of grant number NS41547-01 from NIH/NINDS and Merit Review from VA.

[0002] This application claims priority to U.S. Provisional Patent Application 60/229,406, filed August 31, 2000.

FIELD OF THE INVENTION

[0003] The invention generally relates to the fields of genetics and molecular biology. In particular, the invention relates to the *SCA10* gene and a method of diagnosis of spinocerebellar ataxia type 10.

BACKGROUND OF THE INVENTION

[0004] Autosomal dominant cerebellar ataxias (ADCAs) are currently classified into three groups based on the symptoms presented: ADCA I, ADCA II, and ADCA III. All ADCAs exhibit a degree of cerebellar dysfunction. ADCA I is based on the presence of pyramidal and extrapyramidal symptoms, ophthalmoplegia or a combination thereof. ADCA II is based on the presence of retinopathy. ADCA III is based on the absence of the ADCA I and II symptoms (Zu *et al.*, 1999). A number of SCA genes have been identified or mapped and designated by the HUGO nomenclature committee "spinocerebellar ataxia type n (SCAn)" where "n" refers to the numeric number in order of mapping; these genes include *SCA1*, *SCA2*, *SCA4*, *SCA5*, *SCA6*, *SCA7*, *SCA8*, *SCA10*, *SCA11*, *SCA12*, *SCA13*, *SCA14*, *SCA15*, *SCA16* and *SCA17*. Additionally, two ADCA gene designated *MJD* (also known as *SCA3*) and *DRPLA* have been identified for Machado-Joseph disease and dentatorubral pallidolysian atrophy, respectively. (Silveira *et al.*, 2000). Many of the SCA types are due to a mutation in one of the genes that involves a trinucleotide-repeat expansion in the gene (Zu *et al.*, 1999). The number of repeats often correlates with the age of onset and the severity of the symptoms (O'Sullivan Smith *et al.*, 1999). Diseases other than ADCAs have also been shown to have an expanded trinucleotide repeat as the mutation. Suppression of transcription by a large intronic repeat expansion in Friedreich's ataxia where an expanded GAA repeat interferes with transcription of the *FRDA* gene has been documented

(Bidichandani *et al.*, 1998). Fragile X syndrome and FRAXE mental retardation are caused by CGG and CCG repeat expansion in the 5' untranslated region of the FMR1 and FMR2 gene, respectively. In DM1, an unstable CTG repeat expands up to several thousand copies in the 3' untranslated region of the *DMPK* gene (Wells *et al.*, 1998). Spinocerebellar ataxia type 10 (SCA10) is characterized by gait and limb ataxia, dysarthria, nystagmus, and occasional seizures (Zu *et al.*, 1999; O'Sullivan Smith *et al.*, 1999). SCA10 shows anticipation, in which the onset of the disease is earlier with each passing generation. A pedigree of Mexican-American descent that exhibits spinocerebellar ataxia type 10 has been identified (Zu *et al.*, 1999). SCA10 is believed to be the second most common autosomal dominant cerebellar ataxia after SCA2 in Mexico.

[0005] Spinocerebellar ataxia type 10 (SCA10) is an autosomal dominant disorder characterized by cerebellar ataxia and occasional seizures. Several other SCA subtypes show trinucleotide repeat expansions and prior to the present invention, it was not known whether SCA10 was due to a trinucleotide repeat expansion.

[0006] Genetic mapping studies in two families localized the *SCA10* locus to chromosome 22q13-qter (Matsuura *et al.*, 1999; Zu *et al.*, 1999). Two recombination events narrowed the SCA10 region to a 2.7-cM region between *D22S1140* and *D22S1153* (Matsuura *et al.*, 1999; Zu *et al.*, 1999; Matsuura *et al.*, 1999). Although the DNA sequence of the entire euchromatic part of human chromosome 22 has recently become available, there are still 11 gaps that remain to be sequenced (Dunham *et al.*, 1999). *D22S1160* and *D22S1153* reside in one of these gaps. Nevertheless, two contigs composed of bacterial artificial chromosomes (BACs), phage P1-derived artificial chromosomes (PACs), and cosmids cover most of this region.

[0007] In two large families with SCA10, all patients exhibited pure cerebellar ataxia while 25% and 60% of patients in respective families had recurrent episodes of generalized motor seizures, complex partial seizures, or both (Matsuura *et al.*, 1999; Zu *et al.*, 1999). The clinical phenomenon, known as "anticipation," is common in dominantly inherited SCAs, in which expanded CAG repeats coding for polyglutamine tracts in respective genes are unstable and exhibit larger expansions in successive generations (Orr *et al.*, 1993; Pulst, *et al.*, 1996; Sanpei *et al.*, 1996; Imbert *et al.*, 1996; Kawaguchi *et al.*, 1994; David *et al.*, 1997). Expansion of CAG repeats are involved in SCAs 1, 2, 3, 6, 7, 12, or 17 (Orr *et al.*, 1993; Pulst, *et al.*, 1996; Sanpei *et al.*, 1996; Imbert *et al.*, 1996; Kawaguchi *et al.*,

1994; David *et al.*, 1997; Zhuchenko *et al.*, 1997; Holmes *et al.*, 1999; Nakamura *et al.*, 2001) and expansion of the CTG repeat is found at the *SCA8* locus (Koob *et al.*, 1999).

[0008] Prior to the present invention, the type of and location of the repeat responsible for spinocerebellar ataxia type 10 were unknown. The present invention provides a diagnostic test for spinocerebellar ataxia type 10.

SUMMARY OF THE INVENTION

[0009] One aspect of the present invention a method of detecting spinocerebellar ataxia type 10 in a sample containing DNA from an individual to be tested comprising the step of measuring the presence or absence of DNA expansion at a gene locus associated with spinocerebellar ataxia type 10.

[0010] A further aspect of the present invention is the method of measuring the expansion comprising the steps of: extracting the DNA from a sample to be tested; amplifying the extracted DNA; and identifying the presence or absence of a DNA expansion in the amplified extension products. In an embodiment of the invention, the sample to be tested is selected from the group consisting of blood, semen, vaginal swabs, tissue, mixtures of body fluids, and any biological sample that contains DNA.

[0011] In an embodiment of the invention the amplification is by PCR and may use primers of the sequence of SEQ ID NO: 3 and SEQ ID NO: 4.

[0012] In another embodiment, the DNA expansion is measured by Southern blotting analysis of restriction enzyme digests of genomic DNA with a probe to the *SCA10* locus. The restriction endonuclease may be selected from the group consisting of *EcoRI*, *EcoRV*, *HindIII* and *BglI*. The probe may be created by the use of primers of the sequence of SEQ ID NO: 6 and SEQ ID NO: 7.

[0013] In yet another embodiment, the DNA expansion is determined by pulsed field gel electrophoresis.

[0014] In yet another embodiment, the DNA expansion is determined by fluorescence *in situ* hybridization.

[0015] In an embodiment of the invention, the DNA expansion is comprised of a pentanucleotide repeat. The pentanucleotide repeat may be ATTCT. The number of the pentanucleotide repeat may be between 10 and 29 for unaffected individuals and greater than 800 for individuals affected with spinocerebellar ataxia type 10, with individuals having

pentanucleotide repeats in between the normal and expanded range requiring additional study for a diagnosis of spinocerebellar ataxia type 10. Additional study includes examination for symptoms of spinocerebellar ataxia type 10.

[0016] A further aspect of the invention is a method of detecting pentanucleotide repeats in SCA10 comprising the steps of: isolating DNA from an individual to be tested; and performing PCR analysis using the primers of the sequence of SEQ ID NO: 3 and SEQ ID NO: 4. The pentanucleotide repeat may be ATTCT.

[0017] Another aspect of the invention is a method of diagnosing spinocerebellar ataxia type 10 comprising the steps of: isolating DNA from an individual to be tested; performing PCR analysis using the primers of the sequence of SEQ ID NO: 3 and SEQ ID NO: 4; assessing the number of ATTCT repeats based on comparison to DNA from an unaffected individual; and determining whether the number of ATTCT repeats is expanded in comparison to that of unaffected individuals.

[0018] Another aspect of the present invention is a method of diagnosing spinocerebellar ataxia type 10 comprising the steps of: isolating DNA from an individual to be tested; performing PCR analysis using the primers of the sequence of SEQ ID NO: 10 and SEQ ID NO: 11; and assessing whether the number of ATTCT repeats is expanded in comparison to that of unaffected individuals.

[0019] Another embodiment of the invention is a kit for diagnosis of spinocerebellar ataxia type 10 comprising primers suitable for amplifying *SCA10* and an enzyme suitable for amplifying nucleic acids including various polymerases. The primers may be of the sequence of SEQ ID NO: 3 and SEQ ID NO: 4. In another embodiment, the primers may be of the sequence of SEQ ID NO: 10 and SEQ ID NO: 11.

[0020] Yet another embodiment is a kit for diagnosis of spinocerebellar ataxia type 10 comprising a probe capable of hybridizing to *SCA10* and an enzyme suitable for amplifying nucleic acids including various polymerases to provide the probe to be used in a Southern blot. The probe may be created using primers of the sequence of SEQ ID NO: 6 and SEQ ID NO: 7.

[0021] Another object of the invention is to provide non-human transgenic eukaryotes for use in study of spinocerebellar ataxia 10.

[0022] An embodiment of the invention is a non-human transgenic eukaryote wherein the eukaryote is not expressing *SCA10*; or the ortholog thereof. The eukaryote may be a mammal, and more specifically, may be a mouse.

[0023] Another embodiment of the invention is a non-human transgenic eukaryote that is over-expressing *SCA10*; or the ortholog thereof as compared to a similar non-transgenic eukaryote. The eukaryote may be a mammal, and more specifically, may be a mouse.

[0024] Another embodiment of the invention is a non-human transgenic eukaryote that is expressing the ATTCT repeat array with or without the sequences flanking the array in the *SCA10* gene, as compared to a similar non-transgenic eukaryote. The eukaryote may be a mammal, and more specifically, may be a mouse.

[0025] Other objects, features and advantages of the present invention will become apparent from the following detailed description. It should be understood, however, that the detailed description and the specific examples, while indicating embodiments of the invention, are given by way of illustration only, since various changes and modifications within the spirit and scope of the invention will become apparent to those skilled in the art from this detailed description.

BRIEF SUMMARY OF THE DRAWINGS

[0026] The following drawings form part of the present specification and are included to further demonstrate certain aspects of the present invention. The invention may be better understood by reference to one or more of these drawings in combination with the detailed description of specific embodiments presented herein:

FIG. 1 shows repeat expansion detection (RED) analysis of affected and unaffected individuals.

FIG. 2 shows a Western blot of SCA10 lymphoblastoid cells using mAb 1C2.

FIG. 3a shows pedigrees of four families studied for the SCA10 mutation.

FIG. 3b: shows PCR analysis of the ATTCT pentanucleotide repeat using primers of the sequence of SEQ ID NO:3 and SEQ ID NO:4.

FIG 3c shows Southern analysis of expansion mutations of the ATTCT repeat region.

FIG. 4a shows a schematic presentation of the structure of the *SCA10* gene.

FIG. 4b shows a restriction map of the ATTCT repeat region.

FIG. 4b shows the nucleotide sequence of ATTCT repeat.

FIG. 5 shows distribution of the ATTCT repeat alleles in normal populations.

FIG. 6 shows correlation between the size of expanded SCA10 ATTCT repeat and the age of onset in 26 SCA10 patients.

FIG. 7 shows *in situ* hybridization of radiolabeled probes to horizontal sections of 4-month-old adult (a-d) and 10 day old juvenile (e) mouse brain.

FIG. 8 shows a Northern blot of SCA10 mRNA in lymphoblastoid cell lines of SCA10 patients and unaffected individuals.

FIG. 9 shows PCR analysis of the ATTCT expansion in spinocerebellar ataxia type 10 (SCA10). Lanes 1 and 2 are for SCA10 patients, whereas lanes 3-18 are normal controls.

SEQUENCE SUMMARY

SEQ ID NO: 1. Peptide sequence of human mouse Brain Protein E46-like sequence

SEQ ID NO: 2. Nucleotide sequence of the coding region of human mouse Brain Protein E46-like sequence

SEQ ID NO: 3. Nucleotide sequence of attct-L primer

SEQ ID NO: 4. Nucleotide sequence of attct-R primer

SEQ ID NO: 5. Nucleotide sequence of ATTCT repeat

SEQ ID NO: 6. Nucleotide sequence of DanL primer

SEQ ID NO: 7. Nucleotide sequence of DanR primer

SEQ ID NO: 8. Nucleotide sequence of E46A probe

SEQ ID NO: 9. Nucleotide sequence of E46B probe

SEQ ID NO: 10. Nucleotide sequence of forward primer for PCR

SEQ ID NO: 11. Nucleotide sequence of reverse primer with hanging tail for PCR

SEQ ID NO: 12. Nucleotide sequence of mE46A probe.

SEQ ID NO: 13. Nucleotide sequence for mE46B probe.

DETAILED DESCRIPTION OF THE INVENTION

[0027] The term "a" or "an" as used herein the specification may mean one or more. As used herein in the claim(s), when used in conjunction with the word "comprising", the words "a" or "an" may mean one or more than one. As used herein "another" may mean at least a second or more.

[0028] It will be readily apparent to one skilled in the art that various substitutions and modifications may be made to the invention disclosed herein without departing from the scope and the spirit of the invention.

[0029] As used in this application, the term "polynucleotide having the nucleic acid sequence of SEQ ID NO: 2" refers to a nucleic acid molecule that has been isolated free of total cellular nucleic acid. In some embodiments, the invention concerns a nucleic acid sequence essentially as set forth in SEQ ID NO: 2.

[0030] A functionally equivalent codon is a codon that encodes the same amino acid, such as the six codons for arginine or serine (Table 1), and also refers to codons that encode biologically equivalent amino acids.

TABLE 1

Amino Acids			Codons
Alanine	Ala	A	GCA GCC GCG GCU
Cysteine	Cys	C	UGC UGU
Aspartic acid	Asp	D	GAC GAU
Glutamic acid	Glu	E	GAA GAG
Phenylalanine	Phe	F	UUC UUU
Glycine	Gly	G	GGA GGC GGG GGU
Histidine	His	H	CAC CAU
Isoleucine	Ile	I	AUA AUC AUU
Lysine	Lys	K	AAA AAG
Leucine	Leu	L	UUA UUG CUA CUC CUG CUU
Methionine	Met	M	AUG
Asparagine	Asn	N	AAC AAU

Proline	Pro	P	CCA	CCC	CCG	CCU
Glutamine	Gln	Q	CAA	CAG		
Arginine	Arg	R	AGA	AGG	CGA	CGC
Serine	Ser	S	AGC	AGU	UCA	UCC
Threonine	Thr	T	ACA	ACC	ACG	ACU
Valine	Val	V	GUA	GUC	GUG	GUU
Tryptophan	Trp	W	UGG			
Tyrosine	Tyr	Y	UAC	UAU		

[0031] The DNA segments of the present invention include those encoding biologically functional equivalent SCA10 proteins and peptides, as described above. Such sequences may arise as a consequence of codon redundancy and amino acid functional equivalency that are known to occur naturally within nucleic acid sequences and the proteins thus encoded. Alternatively, functionally equivalent proteins or peptides may be created via the application of recombinant DNA technology, in which changes in the protein structure may be engineered, based on considerations of the properties of the amino acids being exchanged. Changes designed by man may be introduced through the application of site-directed mutagenesis techniques or may be introduced randomly and screened later for the desired function, as described below.

[0032] Naturally, the present invention also encompasses DNA segments that are complementary, or essentially complementary, to the sequence set forth in SEQ ID NO: 2. Nucleic acid sequences that are "complementary" are those that are capable of base-pairing according to the standard Watson-Crick complementary rules. As used herein, the term "complementary" means nucleic acid sequences that are substantially complementary, as may be assessed by the same nucleotide comparison set forth above, or as defined as being capable of hybridizing to the nucleic acid segment of SEQ ID NO: 2 under relatively stringent conditions such as those described herein. Such sequences may encode the entire SCA10 protein or functional or non-functional fragments thereof.

[0033] "DNA expansion" as used herein, refers to an increase in the number of nucleotides found in a particular nucleotide sequence in comparison to that sequence in an individual that is not affected with the condition.

Hybridization

[0034] Alternatively, the hybridizing segments may be shorter oligonucleotides. Sequences of 17 bases long should occur only once in the human genome and, therefore, suffice to specify a unique target sequence. Although shorter oligomers are easier to make and increase *in vivo* accessibility, numerous other factors are involved in determining the specificity of hybridization. Both binding affinity and sequence specificity of an oligonucleotide to its complementary target increases with increasing length. It is contemplated that exemplary oligonucleotides of 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100 or more base pairs will be used, although others are contemplated. Longer polynucleotides encoding 250, 500, 1000, 1250, and 1500 bases and longer are contemplated as well. Such polynucleotides will find use, for example, as probes in Southern and Northern blots and as primers in amplification reactions.

[0035] Accordingly, the nucleotide sequences of the invention may be used for their ability to selectively form duplex molecules with complementary stretches of DNAs and/or RNAs or to provide primers for amplification of DNA or RNA from samples. Depending on the application envisioned, one would desire to employ varying conditions of hybridization to achieve varying degrees of selectivity of the probe or primers for the target sequence.

[0036] In certain applications, for example, substitution of amino acids by site-directed mutagenesis, it is appreciated that lower stringency conditions are required. Under these conditions, hybridization may occur even though the sequences of probe and target strand are not perfectly complementary, but are mismatched at one or more positions. Conditions may be rendered less stringent by increasing salt concentration and decreasing temperature. For example, a medium stringency condition could be provided by about 0.1 to 0.25 M NaCl at temperatures of about 37°C to about 55°C, while a low stringency condition could be provided by about 0.15 M to about 0.9 M salt, at temperatures ranging from about 20°C to about 55°C. Thus, hybridization conditions can be readily manipulated, and thus will generally be a method of choice depending on the desired results.

[0037] In other embodiments, hybridization may be achieved under conditions of, for example, 50 mM Tris-HCl (pH 8.3), 75 mM KCl, 3 mM MgCl₂, 10 mM dithiothreitol, at temperatures between approximately 20°C to about 37°C. Other hybridization conditions

utilized could include approximately 10 mM Tris-HCl (pH 8.3), 50 mM KCl, 1.5 mM MgCl₂, at temperatures ranging from approximately 40°C to about 72°C. Formamide and SDS also may be used to alter the hybridization conditions.

Primers and Probes

[0038] The term primer, as defined herein, is meant to encompass any nucleic acid that is capable of priming the synthesis of a nascent nucleic acid in a template-dependent process. Typically, primers are oligonucleotides from ten to twenty base pairs in length, but longer sequences can be employed. Primers may be provided in double-stranded or single-stranded form, although the single-stranded form is preferred. Probes are defined differently, although they may act as primers. Probes, while perhaps capable of priming, are designed to binding to the target DNA or RNA and need not be used in an amplification process.

[0039] In other embodiments, the probes or primers are labeled with radioactive species (³²P, ¹⁴C, ³⁵S, ³H, or other label), with a fluorophore (rhodamine, fluorescein) or a chemiluminescent (luciferase).

[0040] One method of using probes and primers of the present invention is in the search for genes related to SCA10 or, more particularly, orthologs of SCA10 from other species. Normally, the target DNA will be a genomic or cDNA library, although screening may involve analysis of RNA molecules. By varying the stringency of hybridization, and the region of the probe, different degrees of homology may be discovered.

[0041] In certain embodiments, it will be advantageous to employ nucleic acids of defined sequences of the present invention in combination with an appropriate means, such as a label, for determining hybridization. A wide variety of appropriate indicator means are known in the art, including fluorescent, radioactive, enzymatic or other ligands, such as avidin/biotin, which are capable of being detected. In other embodiments, one may desire to employ a fluorescent label or an enzyme tag such as urease, alkaline phosphatase or peroxidase, instead of radioactive or other environmentally undesirable reagents. In the case of enzyme tags, colorimetric indicator substrates are known that can be employed to provide a detection means that is visibly or spectrophotometrically detectable, to identify specific hybridization with complementary nucleic acid containing samples.

[0042] Another way of exploiting probes and primers of the present invention is in site-directed, or site-specific mutagenesis. Site-specific mutagenesis is a technique useful in the preparation of individual peptides, or biologically functional equivalent proteins

or peptides, through specific mutagenesis of the underlying DNA. The technique further provides a ready ability to prepare and test sequence variants, incorporating one or more of the foregoing considerations, by introducing one or more nucleotide sequence changes into the DNA. Site-specific mutagenesis allows the production of mutants through the use of specific oligonucleotide sequences which encode the DNA sequence of the desired mutation, as well as a sufficient number of adjacent nucleotides, to provide a primer sequence of sufficient size and sequence complexity to form a stable duplex on both sides of the deletion junction being traversed. Typically, a primer of about 17 to 25 nucleotides in length is preferred, with about 5 to 10 residues on both sides of the junction of the sequence being altered.

[0043] In general, it is envisioned that the probes or primers described herein will be useful as reagents in solution hybridization, as in PCR™, for detection of expression of corresponding genes, as well as in embodiments employing a solid phase. Representative solid phase hybridization methods are disclosed in U.S. Patent Nos. 5,843,663, 5,900,481 and 5,919,626. Other methods of hybridization that may be used in the practice of the present invention are disclosed in U.S. Patent Nos. 5,849,481, 5,849,486 and 5,851,772. The relevant portions of these and other references identified in this section of the Specification are incorporated herein by reference.

Template Dependent Amplification Methods

[0044] A number of template dependent processes are available to amplify the marker sequences present in a given template sample. One of the best known amplification methods is the polymerase chain reaction (referred to as PCR) which is described in detail in U.S. Patent Nos. 4,683,195, 4,683,202 and 4,800,159, and in Innis *et al.*, 1990, each of which is incorporated herein by reference in its entirety. Other methods of amplification are ligase chain reaction (LCR), Qbeta Replicase, isothermal amplification, strand displacement amplification (SDA), PCR-like template- and enzyme-dependent synthesis using primers with a capture or detector moiety, transcription-based amplification systems (TAS), cyclical synthesis of single-stranded and double-stranded DNA, "RACE", one-sided PCR, and di-oligonucleotide amplification.

[0045] Briefly, in PCR™, two primer sequences are prepared that are complementary to regions on opposite complementary strands of the marker sequence. An excess of deoxynucleoside triphosphates are added to a reaction mixture along with a DNA polymerase, e.g., Taq polymerase. If the marker sequence is present in a sample, the primers will bind to the marker and the polymerase will cause the primers to be extended along the marker sequence by adding on nucleotides. By raising and lowering the temperature of the reaction mixture, the extended primers will dissociate from the marker to form reaction products, excess primers will bind to the marker and to the reaction products and the process is repeated.

[0046] A reverse transcriptase PCR™ amplification procedure may be performed in order to quantify the amount of mRNA amplified. Methods of reverse transcribing RNA into cDNA are well known and described in Sambrook *et al.*, 1989. Alternative methods for reverse transcription utilize thermostable, RNA-dependent DNA polymerases. These methods are described in WO 90/07641 filed December 21, 1990. Polymerase chain reaction methodologies are well known in the art.

Diagnosing Disorders Involving SCA10

[0047] SCA10 and the corresponding gene can be employed as a diagnostic or prognostic indicator of spinocerebellar ataxia type 10.

Genetic Diagnosis

[0048] One embodiment of the instant invention comprises a method for detecting variation in the number of pentanucleotide repeats in SCA10.

[0049] The methods described herein can be used to detect spinocerebellar ataxia type 10. The method comprises the steps of detecting variation of the (ATTCT)_n repeat found at the 3' end of intron 9 of the *SCA10* gene by measuring the length of the repeat wherein n (number of repeats) for normal is in the range between 10 and 29 and n for spinocerebellar ataxia type 10 is in the range of greater than 800, with individuals having pentanucleotide repeats in between the normal and expanded range requiring additional study for a diagnosis of spinocerebellar ataxia type 10. Additional study includes examination for symptoms of spinocerebellar ataxia type 10.

[0050] The biological sample can be any tissue or fluid. Specifically, the biological sample includes blood, semen, vaginal swabs, buccal swabs, hair follicles, tissue, mixtures of body fluids and any biological sample that contains DNA. Yet other embodiments include dried blood spots, cells of the skin, muscle, fascia, brain, prostate, breast, endometrium, lung, thyroid, parathyroid, adrenal, head and neck, eye, pancreas, small intestine, blood cells, liver, testes, ovaries, colon, skin, stomach, esophagus, spleen, lymph node, bone marrow, thymus or kidney.

[0051] Nucleic acid used is isolated from cells contained in the biological sample, according to standard methodologies (Sambrook *et al.*, 1989). The nucleic acid may be genomic DNA or fractionated or whole cell RNA. Where RNA is used, it may be desired to convert the RNA to a complementary DNA. In one embodiment, the RNA is whole cell RNA and/or polyA-selected RNA. Normally, the nucleic acid is amplified.

[0052] Depending on the format, the specific nucleic acids of interest are identified in the sample directly using amplification or with a second, known nucleic acid following amplification. Next, the identified product is detected. In certain applications, the detection may be performed by visual means (e.g., ethidium bromide staining of a gel). Alternatively, the detection may involve indirect identification of the product via chemiluminescence, radioactive scintigraphy of radiolabel or fluorescent label or even via a system using electrical or thermal impulse signals (Affymax Technology; Bellus, 1994).

[0053] Following detection, one may compare the results seen in a given patient with a statistically significant reference group of normal patients. In this way, it is possible to correlate the amount of repeats detected with various clinical states.

[0054] Alterations of a gene include deletions, insertions, point mutations and duplications. Point mutations result in stop codons, frameshift mutations or amino acid substitutions. Somatic mutations are those occurring in non-germline tissues. Germ-line tissue can occur in any tissue and are inherited. Mutations in and outside the coding region also may affect the amount of SCA10 produced, both by altering the transcription of the gene or in destabilizing or otherwise altering the processing of either the transcript (mRNA) or protein.

[0055] A variety of different assays are contemplated in this regard, including but not limited to, PCR, fluorescent *in situ* hybridization (FISH), direct DNA sequencing, PFGE analysis, Southern blotting, and RFLP.

10055977.1

[0056] As used herein the term "polymerase chain reaction" or "PCR" refers to the PCR procedure described in the patents to Mullis, *et al.*, U.S. Patent Nos. 4,683,195 and 4,683,202. The procedure basically involves: (1)treating extracted DNA to form single-stranded complementary strands; (2)adding a pair of oligonucleotide primers, wherein one primer of the pair is substantially complementary to part of the sequence in the sense strand and the other primer of each pair is substantially complementary to a different part of the same sequence in the complementary antisense strand; (3)annealing the paired primers to the complementary sequence; (4)simultaneously extending the annealed primers from a 3' terminus of each primer to synthesize an extension product complementary to the strands annealed to each primer wherein said extension products after separation from the complement serve as templates for the synthesis of an extension product for the other primer of each pair; (5)separating said extension products from said templates to produce single-stranded molecules; and (6)amplifying said single-stranded molecules by repeating at least once said annealing, extending and separating steps.

[0057] As used herein fluorescence *in situ* hybridization or "FISH" refers to the procedure described in Wott *et al.* (1988) and Kievits *et al.* (1990). The procedure basically involves the steps of preparing interphase or metaphase spreads from cells of peripheral blood lymphocytes and hybridizing labeled probes to the interphase or metaphase spreads. Using probes with mixed labels allows visualization of space, order and distance between hybridization sites. After hybridization the labels are examined to determine the order and distance between the hybridization sites.

[0058] As used herein, the term "pulsed field gel electrophoresis" or "PFGE" refers to a procedure described by Schwart *et al.* (1982). The procedure basically comprises running a standard electrophoresis gel (agarose, polyacrylamide or other gel known to those skilled in the art) under pulsing conditions. One skilled in the art recognizes that the strength of the field as well the direction of the field is pulsed and rotated in order to separate megabase DNA molecules. Current commercial systems are computer controlled and select the strength, direction and time of pulse depending on the molecular weight of DNA to be separated.

[0059] Measuring the amount or length of the repeat can be done by using FISH. In the FISH method, the repetitive sequence can be used as a probe to distinguish between normal and spinocerebellar ataxia type 10 syndrome simply by the presence or

absence of a signal to the repetitive sequence. In this case, the application of the repeat sequence provides a sufficiently large target for the hybridization. Thus, it is possible that very sensitive FISH might detect affected individuals even though these would be lost to routine microscopy and detection. Although FISH is usually applied to metaphase nuclei, in the present invention it is applicable to both metaphase and interphase for the detection of spinocerebellar ataxia type 10.

[0060] Alternate methods to measure the repeat can include visual examination, densitometry measurement, quantitative radioactivity and quantitative fluorescence.

Southern/Northern Blotting

[0061] In one embodiment the size of the repeat is determined by Southern blotting analysis of restriction enzyme digests with probes contained within the SCA10 gene region.

[0062] One embodiment of the present invention is a method of detecting spinocerebellar ataxia type 10 syndrome comprising the steps of digesting DNA from an individual to be tested with a restriction endonuclease and detecting the restriction fragment length to polymorphism (RFLP) with hybridization to probes within the spinocerebellar ataxia type 10

[0063] Blotting techniques are well known to those of skill in the art. Southern blotting involves the use of DNA as a target, whereas Northern blotting involves the use of RNA as a target. Each provide different types of information, although cDNA blotting is analogous, in many aspects, to blotting of RNA species.

[0064] Briefly, a probe is used to target a DNA or RNA species that has been immobilized on a suitable matrix, often a filter of nitrocellulose. The different species should be spatially separated to facilitate analysis. This often is accomplished by gel electrophoresis of nucleic acid species followed by "blotting" on to the filter.

[0065] Subsequently, the blotted target is incubated with a probe (usually labeled) under conditions that promote denaturation and rehybridization. Because the probe is designed to base pair with the target, the probe will binding a portion of the target sequence under renaturing conditions. Unbound probe is then removed, and detection is accomplished as described above.

Separation Methods

[0066] It normally is desirable, at one stage or another, to separate the amplification product from the template and the excess primer for the purpose of determining whether specific amplification has occurred. In one embodiment, amplification products are separated by agarose, agarose-acrylamide or polyacrylamide gel electrophoresis using standard methods. See Sambrook *et al.*, 1989.

[0067] Alternatively, chromatographic techniques may be employed to effect separation. There are many kinds of chromatography which may be used in the present invention: adsorption, partition, ion-exchange and molecular sieve, and many specialized techniques for using them including column, paper, thin-layer and gas chromatography (Freifelder, 1982).

Detection Methods

[0068] Products may be visualized in order to confirm amplification of the marker sequences. One typical visualization method involves staining of a gel with ethidium bromide and visualization under UV light. Alternatively, if the amplification products are integrally labeled with radio- or fluorometrically-labeled nucleotides, the amplification products can then be exposed to x-ray film or visualized under the appropriate stimulating spectra, following separation.

Kit Components

[0069] All the essential materials and reagents required for detecting and sequencing *SCA10* and variants thereof may be assembled together in a kit. This generally will comprise preselected primers and probes. Also included may be enzymes suitable for amplifying nucleic acids including various polymerases (RT, Taq, SequenaseTM etc.), deoxynucleotides and buffers to provide the necessary reaction mixture for amplification. Such kits also generally will comprise, in suitable means, distinct containers for each individual reagent and enzyme as well as for each primer or probe.

Transgenic Animals/Knockout Animals

[0070] In one embodiment of the invention, transgenic animals are produced which contain a functional transgene encoding a functional SCA10 polypeptide or variants thereof, or a transgene that is transcribed into an RNA but not translated into a protein. Transgenic animals expressing *SCA10* transgenes can be used as models for studying spinocerebellar ataxia type 10.

[0071] In one embodiment of the invention, a SCA10 transgene is introduced into a non-human host to produce a transgenic animal expressing a human or murine SCA10 gene. The transgenic animal is produced by the integration of the transgene into the genome in a manner that permits the expression of the transgene. Methods for producing transgenic animals are generally described by Wagner and Hoppe (U.S. Patent No. 4,873,191; which is incorporated herein by reference), Brinster *et al.* 1985; which is incorporated herein by reference in its entirety) and in "Manipulating the Mouse Embryo; A Laboratory Manual", 1994; which is incorporated herein by reference in its entirety).

[0072] It may be desirable to replace the endogenous SCA10 by homologous recombination between the transgene and the endogenous gene; or the endogenous gene may be eliminated by deletion as in the preparation of "knock-out" animals. Typically, a *SCA10* gene flanked by genomic sequences is transferred by microinjection into a fertilized egg. The microinjected eggs are implanted into a host female, and the progeny are screened for the expression of the transgene. Transgenic animals are produced from the fertilized eggs from a number of animals including, but not limited to reptiles, amphibians, birds, mammals, and fish. Within a particular embodiment, transgenic mice are generated which overexpress *SCA10* or express a mutant form of the polypeptide. Alternatively, the absence of a SCA10 in "knock-out" mice permits the study of the effects that loss of SCA10 protein has on a cell *in vivo*.

[0073] Promoter sequences may be used to drive expression of β -galactosidase or other reporter sequences such as green fluorescent protein, chloramphenicol acetyl transferase, and others known in the art.. A reporter construct in transgenic mice is used to identify factors which regulate SCA10 expression.

[0074] One embodiment of the present invention is a method of detecting spinocerebellar ataxia type 10 comprising the step of measuring the presence or absence of DNA expansion at a gene locus associated with spinocerebellar ataxia type 10.

[0075] An embodiment of the present invention is the method of measuring the expansion comprising the steps of extracting the DNA from a sample to be tested, amplifying the extracted DNA, and identifying the presence or absence of a DNA expansion in the amplified extension products. The presence of a DNA expansion over the sequence found in unaffected individuals indicates the sample came from an individual with spinocerebellar ataxia. Suitable samples for testing include blood, semen, vaginal swabs, tissue, mixtures of body fluids, and any biological sample that contains DNA.

[0076] In an embodiment of the invention the amplification of the extracted DNA is by PCR and the primers are of the sequence of SEQ ID NO: 3 and SEQ ID NO: 4.

[0077] In another embodiment, the DNA expansion is measured by Southern blotting analysis of restriction enzyme digests with a probe to the *SCA10* locus. The digest may be performed using the restriction endonucleases EcoRI, EcoRV, HindIII and BglI. The probe may be created by the use of primers of the sequence of SEQ ID NO: 6 and SEQ ID NO: 7.

[0078] In other embodiments, the DNA expansion is determined by pulsed field gel electrophoresis or fluorescence *in situ* hybridization.

[0079] The DNA expansion is comprised of a pentanucleotide repeat. The pentanucleotide repeat is of the sequence ATTCT and the length of the pentanucleotide repeat is between 10 and 29 for unaffected individuals and greater than 800 for individuals affected with spinocerebellar ataxia type 10, with individuals having pentanucleotide repeats in between the normal and expanded range requiring additional study for a diagnosis of spinocerebellar ataxia type 10. Additional study includes examination for symptoms of spinocerebellar ataxia type 10.

[0080] Another embodiment of the invention is a kit for diagnosis of spinocerebellar ataxia type 10 comprising primers suitable for amplifying SCA10 and an enzyme suitable for amplifying nucleic acids including various polymerases. The primers may be of the sequence of SEQ ID NO: 3 and SEQ ID NO: 4.

[0081] Another embodiment is a kit for diagnosis of spinocerebellar ataxia type 10 comprising a probe capable of hybridizing to SCA10 and an enzyme suitable for amplifying nucleic acids including various polymerases to provide the probe to be used in a Southern blot. The probe may be created using primers of the sequence of SEQ ID NO: 6 and SEQ ID NO: 7.

[0082] Another object of the invention is to provide non-human transgenic eukaryotes for use in study of spinocerebellar ataxia 10.

[0083] Another embodiments of the invention is a non-human transgenic eukaryote wherein the eukaryote is not expressing SCA10; or the ortholog thereof. Yet another embodiment is a non-human transgenic eukaryote that is over-expressing SCA10; or the ortholog thereof as compared to a similar non-transgenic eukaryote.

Examples

[0084] The following examples are included to demonstrate preferred embodiments of the invention. It should be appreciated by those of skilled the art that the techniques disclosed in the examples which follow represent techniques discovered by the inventor to function well in the practice of the invention, and thus can be considered to constitute preferred modes for its practice. However, those of skill in the art should, in light of the present disclosure, appreciate that many changes can be made in the specific embodiments which are disclosed and still obtain a like or similar result without departing from the concept, spirit and scope of the invention. More specifically, it will be apparent that certain agents which are both chemically and physiologically related may be substituted for the agents described herein while the same or similar results would be achieved. All such similar substitutes and modifications apparent to those skilled in the art are deemed to be within the spirit, scope and concept of the invention as defined by the appended claims. In the examples all percentages are by weight, if for solids and by volumes, if for liquids and all temperatures are in degrees Celsius unless otherwise noted.

Example 1

DNA extraction and polymorphism analysis

[0085] DNA was extracted from blood samples of all participants. Recombination events were defined and haplotypes were constructed by analyzing genomic DNA by PCR amplification of polymorphic simple sequence repeat (SSR) markers using end-labeled primers. The size of the PCR products were determined on 6% denaturing polyacrylamide gels.

Example 2

Repeat expansion detection (RED) analysis

[0086] Because of the dominant inheritance of SCA10 and presence of anticipation, 14 trinucleotide repeats (>3 repeats in length) listed in the SCA10 candidate interval in the chromosome 22 genome database at the Sanger Centre (Dunham *et al.*, 1999) were examined. None of these repeats showed expansions in the affected members of the families.

[0087] RED analysis (Schalling *et al.*, 1993, Koob *et al.*, 1998) was performed using genomic DNA samples of SCA10 patients and control subjects devoid of CAG/CTG expansions at the *ERDA1* (chromosome 17) (Ikeuchi *et al.*, 1998) and *SEF2.1* (chromosome 18) (Breschel *et al.*, 1997) loci. After 400 cycles of denaturing and annealing (95 °C for 10 s; 65 °C for 30 s) of 10 µg genomic DNA and (CTG)₁₀ or (TTG)₁₀ oligonucleotides in the presence of 5 U Ampligase (Epicentre), the product underwent electrophoresis and was then transferred to a nylon membrane. The membrane was hybridized with a ³²P-end-labeled (CAG)₁₀ or (CAA)₁₀ oligonucleotide probe to detect tandemly ligated (CTG)₁₀ or (TTG)₁₀ oligonucleotides (FIG. 1). Repeat expansion detection (RED) analysis of genomic DNA samples from affected family members failed to show evidence of a CAG or CAA expansion. There was no difference between the affected (70, 40 repeats) and non-affected (40, 70 repeats).

Example 3

Western blot analysis

[0088] Protein extracts obtained from lymphoblastoid cell lines derived from SCA10 patients and control subjects were electrophoresed. After western blotting, the monoclonal antibody "1C2" raised against TATA-binding protein (Trottier *et al.*, 1995) was used to look for expanded polyglutamine repeats.

[0089] Proteins containing expanded polyglutamine tracts were not detected (FIG. 2). No detection of abnormal proteins in SCA10 lymphoblastoid cells (LCs) were

found compared with the control. The mAb 1C2 raised against TATA-binding protein (TBP) detected ataxin-1 at 100K in SCA1 LCs with a large CAG expansion (n=82) as well as TBP.

Example 4

PCR assay of SCA10 repeats

[0090] The ATTCT repeat region from genomic DNA of affected and unaffected individuals was PCR-amplified using primers of the sequence (5'-AGAAAACAGATGGCAGAATGA-3') (SEQ ID NO:3) and (5'-GCCTGGGCAACATAGAGAGA-3') (SEQ ID NO:4) in HotStarTaq Master Mix (Qiagen) including 10% DMSO (FIG. 4c). The PCR condition consisted of initial denaturing at 95 °C for 15 min, 30 PCR cycles (denaturing at 94 °C for 30 s, annealing at 54 °C for 30 s, extension at 72 °C for 45 s), and additional extension at 72 °C for 7 min. The PCR products from affected and unaffected individuals were sequenced using Applied Biosystems (ABI) automated sequencer with fluorescent dideoxynucleotides.

[0091] The ATTCT pentanucleotide repeat region was PCR-amplified from the genomic DNA samples of the family members indicated in FIG. 3a using primers of the sequence of SEQ ID NO: 3 (attct-L) and SEQ ID NO: 4 (attct-R). All affected individuals showed a single allele of variable size in families a, b, c and d (note that each band accompanies a shadow band underneath due to PCR artifact). In family a, two unaffected individuals (I-1 and III-2) are heterozygous and two spouses (II-1, III-5) are homozygous for the ATTCT repeat. In this family, affected individuals in the second generation (II-2 and II-3) failed to transmit their 12-repeat allele to their affected offspring (III-1, III-3, III-4, III-6, III-7, III-8 and III-9) while an unaffected offspring (III-2) received this allele from the affected father (II-2). The alleles of unaffected parents (I-1 and II-1) were passed on to their offspring in a pattern consistent with Mendelian inheritance. These data indicate that the affected individuals are apparently hemizygous for the ATTCT repeat.

[0092] PCR analysis of the region spanning a pentanucleotide (ATTCT) tandem repeat in intron 9 of the *SCA10* gene (previously designated *E46*) (FIG. 4) showed repeat number polymorphism in normal individuals (FIG. 3b, FIG. 5). Normal alleles contained from 10 to 20 repeats with 82.1% heterozygosity. Sequence analysis of the alleles obtained from 20 normal individuals showed tandem repeats of ATTCT without interruption.

The distribution of the alleles was unimodal with similar patterns among the Caucasian, Mexican and Japanese populations examined (FIG. 5). The allele distributions in each of three ethnically defined populations (including 127 persons from the Mexican population) were consistent with the Hardy-Weinberg equilibrium ($p \gg 0.05$). In SCA10 families, PCR analysis demonstrated a uniform lack of heterozygosity in all affected individuals and carriers of the disease haplotype (Matsuura *et al.*, 1999, Zu *et al.*, 1999), with the single amplified allele of the ATTCT repeat being shared by their unaffected parent. When the affected and unaffected parents of the patients carried distinct ATTCT repeat alleles, the single allele amplified from the affected parent was never transmitted to any of these patients, indicating that the affected parent was hemizygous and that only the allele on the wild-type (non-SCA10) chromosome is amplified (FIG. 3b). Two other sets of primers flanking the ATTCT repeat gave the same result, excluding the possibility that the apparent hemizygosity is due to a mutation within the primer binding site.

Example 5

Southern analysis

[0093] Southern blots of genomic DNA digested with multiple restriction enzymes, including *EcoRI*, *EcoRV*, *HindIII* and *BglII*, underwent hybridization to a non-repetitive probe obtained by PCR amplification of the region immediately upstream of the ATTCT repeat using DNA from a PAC clone (RP1-37M3) as template (FIG. 4b).

[0094] Ten μ g of *EcoRI*-digested genomic DNA underwent 0.6 % to 0.8 % agarose gel electrophoresis followed by capillary transfer to the Hybond N⁺ membrane (Amersham). An 800-bp SCA10 probe located upstream to the pentanucleotide repeat was generated by PCR using DNA of a PAC clone (RP1-37M3) as the template (FIG. 4b). Primers were: DanL (5'-TCCTTCCTCAGTCTTTCTGG-3') (SEQ ID NO: 6) and DanR (5'-TGCCATCTGTTTTCTATTTG-3') (SEQ ID NO: 7). Using the probe random prime labeled with ³²P- α -dCTP (Amersham), the membrane was hybridized in Church buffer (0.1 mM EDTA at pH 8.0, 0.5 M sodium phosphate at pH 7.2, 7% SDS) at 60 °C overnight and analyzed by autoradiography after washing.

[0095] Southern blots of the genomic DNA samples digested with *EcoRI* using a 0.6 % to 0.8 % agarose gel show variably expanded alleles in affected members of the

families shown above. All individual examined have a normal allele (2.5 kb). The apparent variability of the normal allele size is attributable to gel-loading artifacts since additional analyses using the same (*EcoRI*) and different (*EcoRV*, *HindIII* and *BglI*) restriction enzymes did not show consistent variability of the normal allele size. The genotype of each individual is shown at the bottom, with an estimated number of pentanucleotide repeats for disease-chromosomes based on the fragment size.

[0096] In addition to the expected normal allele, a variably expanded allele in all affected individuals was detected, while in all unaffected family members only the wild-type allele was detected, demonstrating that the ATTCT repeat region is expanded exclusively in SCA10 patients (FIG. 3c).

Example 6

Pentanucleotide repeat expansions

[0097] Three additional families with an autosomal dominant disease characterized by cerebellar ataxia and seizures were identified (FIG. 3a). The age of disease onset appeared to be earlier in successive generations in these families. This clinical phenomenon is known as "anticipation."

[0098] The inventors found an expansion of a pentanucleotide (ATTCT) repeat in intron 9 of the *SCA10* gene, up to 19 kb larger than the normal allele, in all patients in five SCA10 families. Analysis of 562 chromosomes from unaffected individuals of various ethnic origins, showed a size range of 10 to 29 ATTCT repeats with no evidence of expansions. The data indicates that the novel SCA10 intronic ATTCT pentanucleotide repeat in SCA10 patients is unstable.

[0099] An SCA10 expansion mutation was not detected in 17 probands of unrelated Caucasian-American families with autosomal dominant cerebellar ataxia that have tested negative for the SCAs 1, 2, 3, 6, 7, 8 or 12 mutations.

Example 7

Physical map of the ATTCT pentanucleotide repeat region.

a: A schematic presentation of the structure of the *SCA10* gene.

[0100] *SCA10* consists of 12 exons which total 1971 base pairs. The ATTCT repeat is located in intron 9. The 12 exons of *SCA10* span 172.8 kb of genomic DNA with an open reading frame (ORF) of 1428 bp, encoding 475 amino acids with no homology to known human proteins. Although human *SCA10* is well-conserved with its presumed mouse ortholog (82% identity, 91% similarity over 475 amino acids), the next most similar sequence present in the GenBank database is a putative plant protein of unknown function identified by the *Arabidopsis* genome project (24% identical, 41% similar over 409 amino acids). Analysis of the *SCA10* amino acid sequence indicates that this protein is a non-transmembrane globular protein without any nuclear localization signal or other type of signal peptide (Golgi, peroxisomal, vacuolar, or endoplasmic reticulum-retention). The gap at the left of PAC 37M3 does not represent missing sequence, but was introduced to preserve scale.

b: A restriction map of the ATTCT repeat region defined by flanking *HindIII* restriction sites (nt 17,023 and 34,567).

[0101] The numbers are nucleotide positions in the PAC37M3. "Probe" indicates the position of the probe used (nt 25,222-26,021) to detect the 2.5 kb *EcoRI* fragment shown in FIG. 4b in the Southern analysis. The ATTCT repeat is located downstream of the probe within the 2.5 kb *EcoRI* fragment.

c: Nucleotide sequence of the ATTCT repeat (14 repeats; underlined, nt 26,101-26170) and the flanking regions (nt 25,981-26,281).

[0102] Arrows underline PCR primer sequences (attct-L, SEQ ID NO: 3) and (attct-R, SEQ ID NO: 4) that were used for amplification of the ATTCT repeat region shown in Figure 4c.

Example 8

Correlations

[0103] The variable size of expanded alleles among affected individuals indicates the expanded ATTCT repeat is highly unstable. In the SCA10 families studied, intergenerational changes of the expanded alleles were variable, including not only expansions but also contractions (FIG. 3c). Haplotype analysis using DNA markers flanking the ATTCT repeat showed a single disease haplotype shared by all SCA10 families studied. A potentially important observation in the SCA10 families was the weak inverse correlation between the size of the expanded allele and the age of disease onset ($n = 26$, $r^2 = 0.34$, $p = 0.018$) (FIG. 6). The correlation argues against the possibility that the expansion represents a rare non-pathogenic polymorphism. Furthermore, there is no evidence of the ATTCT expansion in 562 normal chromosomes.

Example 9

Distribution of the ATTCT repeat alleles in normal populations.

[0104] Shown is a histogram of the normal ATTCT repeat alleles in Caucasian ($n = 250$), Japanese ($n = 100$) and Mexican ($n = 254$) chromosomes (FIG. 5). The distribution of the alleles was unimodal with similar patterns among the Caucasian, Mexican and Japanese populations examined. There were, however, subtle differences between these populations. In the Mexican and Japanese populations, the 14-ATTCT-repeat allele was the modal allele with frequencies of 29% and 27%, respectively. In the Caucasian population, the 13-repeat allele was modal with a frequency of 30%.

Example 10

Correlation between the size of expanded SCA10 ATTCT repeat and the age of onset.

[0105] A scatter plot (FIG. 6) shows an inverse correlation between the size of expansion and the age of onset in 26 SCA10 patients ($r^2 = 0.34$, $p = 0.018$). Each symbol represents an SCA10 patient, and the linear regression line is shown.

Example 11

In-situ hybridization

[0106] Horizontal brain sections (12 μ m) were fixed from male C57BL/6J mice in 4% paraformaldehyde in PBS followed by dehydration. Antisense oligonucleotide probes were end-labeled using terminal deoxynucleotidyl transferase (Promega) and [α -³⁵S] dATP (1250 Ci/mmol; NEN) to a specific activity of $\sim 10^9$ dpm/ μ g. Hybridization solution contained 50% (v/v) Formamide, 4X SSC, 25mM sodium phosphate, 1mM sodium pyrophosphate, 10% dextran sulfate (w/v), 5X Denhardt's solution, 200 μ g/ml sonicated herring sperm DNA (Promega), 100 μ g/ml polyadenylic acid [5'] (Sigma-Aldrich) and 5×10^2 dpm of [α -³⁵S] dATP-labeled probe. To monitor background hybridization, the inventors added 100-fold excess of unlabeled oligonucleotide to control sections. The sequence of the 45-mer probes was:

	mE46A	5'-	
CTGTTGTCTTCAGTGAGATTTTCGCACAGCATACACCACCCACTGC-3'	(SEQ ID		
NO:12)	and	mE46B	5'-
	CACTGCAGAGATGAGAGGTCCGTGAGATGGAATCTGAATGTGTTC-3'	(SEQ ID	
	NO:13).		

Sections were hybridized to probes overnight at 42 °C, washed in 1X SSC (22 °C, 20 min), 0.3X SSC (55 °C, 40 min), and 2X SSC (22 °C, 5 min) then dehydrated and exposed to Kodak BioMax MR film for 1 week.

[0107] The *in situ* hybridization was to horizontal sections of 4-month-old adult (a-d) and 10 day old juvenile (e) mouse brain (FIG. 7). Expression was similar to the pattern of cell density determined by cresyl violet staining of the same sections. A-d: dorsal to ventral progression; f: negative control for non-specific hybridization to an adult brain section.

Example 12

SCA10 Expression

[0108] SCA10 is expressed widely in mammalian brain (FIG. 7), consistent with the phenotype of ataxia and epilepsy, and expressed sequence tags representing SCA10

are also detectable in several non-neuronal tissues. The northern blot data show no obvious changes in the level of SCA10 mRNA in lymphoblastoid cell lines of SCA10 patients (FIG. 8).

Example 13

PCR amplification of expanded ATTCT repeat in SCA10

[0109] Unlike most ADCAs with triplet repeat (CAG/CTG) expansions, the mutation of *SCA10* is an unstable expansion of an ATTCT pentanucleotide repeat (800 to about 4500 repeats; normal range 10-29 repeats). Because the expanded ATTCT repeat can give alleles of greater than 20 kb, which may be too large to be amplified by PCR, the molecular diagnosis of SCA10 can consist of two steps: 1) PCR analysis of the region spanning the ATTCT repeat to analyze the size of the normal alleles, and 2) Southern blot analysis to determine whether the ATTCT repeat is expanded in individuals showing a single allele by PCR. However, this method is laborious and costly and is difficult to use with degraded DNA samples or samples with a limited amount of DNA.

[0110] One aspect of the present invention is a simple PCR system that detects the expanded ATTCT allele with a characteristic ladder. The reaction was performed in 10 μ l HotStarTaq Master Mix (Qiagen) containing 200 ng of genomic DNA as a template and 1 μ M each of the primers. The ³²P-end-labelled forward primer (GAAGACAAATAGAAAACAGATGGCAGA) (SEQ ID NO: 10) corresponds to a unique sequence upstream of the ATTCT repeat, while the reverse primer consists of a repeat sequence with a hanging tail sequence at its 5' end (TACGCATCCCAGTTTGAGACGG(AATAG)₈) (SEQ ID NO: 11) (Warner *et al.*, 1996). The 5' hanging sequence has negligible complementarity to itself, AATAG or CTATT repeats, or any known human sequences. The repeat region at the 3' terminus of SEQ ID NO: 11 should randomly bind at multiple sites within the ATTCT repeat tract, generating a mixture of products that contain a variable number of repeats during the first PCR cycle. The hanging tail at the 5' end of the PCR products serve as an anchor that increases the probability for the reverse primer to anneal at the end of the repeat sequence during the subsequent cycles, preventing progressive shortening of the successive PCR products. The

PCR conditions consisted of an initial denaturing at 95 °C for 15 min, 30 PCR cycles (94 °C for 30 s, 60 °C for 30 s, and 72 °C for 2 min), and additional extension at 72 °C for 10 min. PCR products were electrophoresed on 6% denaturing polyacrylamide gels and visualized by autoradiography.

[0111] Analysis were performed on 44 SCA10 patients and 8 SCA10 mutation carriers confirmed by Southern blot and 100 normal controls whose ATTCT repeat alleles had been determined. Results of some of these analysis are shown in Figure 9. All SCA10 patients showed continuous ladder exceeding the product range shown by normal controls. The repeat number of the largest product in the ladder corresponded to the known size of the larger allele in each normal control. There were no false positive results, i.e., no expansions in the normal controls. Moreover, expanded alleles were detected in highly degraded DNA samples of the three affecteds in two SCA10 pedigrees where Southern analysis failed to give a reliable result.

[0112] This analysis quickly and accurately identifies the mutation status of ATTCT repeat alleles. It easily distinguishes the individuals homozygous for normal alleles from SCA10 patients and reduces the number of samples that require the Southern blot analysis. This method is also useful for the cloning of the expanded alleles, which previously has been unsuccessful due to the massive size of the expansion. Uses of the cloned expanded repeats are for studies of the repeat structure and construction of the transgene containing the expanded ATTCT repeats. This PCR-based technology is useful for both genetic diagnosis and investigation of the disease mechanisms in SCA10.

SEQUENCE LISTING

<110> Ashizawa, Tetsuo

Tohru, Matsuura

<120> DNA TEST FOR SCA-10

<130> P02039US1/10023139/OTA01-01

<140> TBA

<141> 2001-08-29

<150> US 60/229,406

<151> 2000-08-29

<160> 13

<170> PatentIn version 3.0

<210> 1

<211> 475

<212> PRT

<213> * HUMAN

<400> 1

Met Ala Ala Pro Arg Pro Pro Pro Ala Arg Leu Ser Gly Val Met Val
1 5 10 15

Pro Ala Pro Ile Gln Asp Leu Glu Ala Leu Arg Ala Leu Thr Ala Leu
20 25 30

Phe Lys Glu Gln Arg Asn Arg Glu Thr Ala Pro Arg Thr Ile Phe Gln

[illegible]

Asn Cys Gly Cys Val Arg Ala Glu Gly Asp Ile Ser Asn Val Ala Asn
355 360 365

Gly Phe Lys Ser His Leu Ile Arg Leu Ile Gly Asn Leu Cys Tyr Lys
370 375 380

Asn Lys Asp Asn Gln Asp Lys Val Asn Glu Leu Asp Gly Ile Pro Leu
385 390 395 400

Ile Leu Asp Asn Cys Asn Ile Ser Asp Ser Asn Pro Phe Leu Thr Gln
405 410 415

Trp Val Ile Tyr Ala Ile Arg Asn Leu Thr Glu Asp Asn Ser Gln Asn
420 425 430

Gln Asp Leu Ile Ala Lys Met Glu Glu Gln Gly Leu Ala Asp Ala Ser
435 440 445

Leu Leu Lys Lys Val Gly Phe Glu Val Glu Lys Lys Gly Glu Lys Leu
450 455 460

Ile Leu Lys Ser Thr Arg Asp Thr Pro Lys Pro
465 470 475

<210> 2

<211> 1971

<212> DNA

<213> HUMAN

<400> 2

```

ggctcagcct agagctctcc ggcggcgggc cagcttcagg gcagcgcggg ctgcagcggc      60
ggcggcggtt agggctgtgt agggcgaggc ctcccccttc ctctctgcca tcctactcct      120
ccctcctcgt catctctccc ctctgtcttc ctgccttcc tcctcctcgt caggctcgac      180
ccagctgtga gcggcaagat ggcggcgccc aggccgcgcg ctgccaggct gtcgggcgtc      240
atggtgcggg cgcccatcca agacctggag gccctgcgcg cgctcacggc gctcttcaaa      300
gagcagcggg accgagaaac agcaccagg actatcttcc aaagagttct ggatataccta      360
aagaaatctt ctcatgctgt tgagcttgcc tgcagagatc catcccaagt ggaaaacctg      420
gcttccagtc tgcagttaat aacagaatgc ttcagggtgc ttcgcaatgc ttgcatagag      480
tgttctgtga accagaattc aatcaggaac ttggatacga ttggtgttgc tgttgatttg      540
attcttctgt ttcgtgaact gcgagtggaa caggaatctc tgttgacagc ttttcgctgt      600
ggcctgcagt ttttaggcaa cattgcctca cggaatgaag attccagtc tattgtttgg      660

```


gtgcatgctt tcccagaact gtttttgtct tgcttaaadc atccggacaa aaaaattgtt 720
 gcctactcct caatgatttt gtttacatcc cttaatcatg aaagaatgaa agaactggag 780
 gagaacctca atattgcaat tgatgtcata gatgcttacc aaaaacatcc tgaatcagaa 840
 tggcogttct tgattattac agacctcttt ctgaaaagcc cggaattggt acaagccatg 900
 tttoccaaac tgaacaatca agaaagagtt acactgttag accttatgat agccaagata 960
 acgagtgatg agccactcac caaggatgac atccctgtgt ttttgcgga tgctgagttg 1020
 attgcaagca cctttgtgga tcagtgcag actgtgctca agctggcctc tgaggagcct 1080
 cctgatgatg aggaggcact ggctacaatt aggcttctcg acgtcctgtg cgaaatgact 1140
 gtgaatactg agctgctcgg ctatctgcag gttttccctg gcttgctgga aagagtgatt 1200
 gatcttttgc ggggtgattca thtagctgga aaagaaacca caaacatctt cagtaattgt 1260
 gggtgcgtga gagcagaagg tgacatctcc aatgtggcca atgggtttta gtctcatctc 1320
 attcgtctga ttggaaatct gtgttacaag aataaagata accaagacaa ggtaaagtga 1380
 ctggatggta tcccgttgat cctggacaac tgcaacatca gtgacagtaa cccctttctg 1440
 acccagtggg tgatatatgc catccgaaac cttaccgaag acaacagcca aaaccaagat 1500
 ttgattgcaa agatggagga acaggggctg gcagatgcat ccctacttaa aaaagtgggt 1560
 tttgaagttg aaaagaaagg cgaaaagctg atcctgaaat ctactagaga cccccctaag 1620
 ccatgaatga actacatcca aatacctgaa tttttggaat ctgtttcatg gatttttcat 1680
 cttctaccgt atgtgaaatt gcaagtgttt gaagatttat aagtacaaat ttgggaacat 1740
 acaaatcttt taggtagtag agtttaacgt gtataagcta aaagtgaag taactgagt 1800
 ttctcttgtt tctttgcatt aatgtaactg tgtggtttgc ctttgtcccc ctggatagaa 1860
 cgtgcattta aagaatatat tgtacttact gtgacagcag ataataaacc agtctcttgg 1920
 agggcaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa a 1971

<210> 3

<211> 21

<212> DNA

<213> PRIMERS

<220>

<221> misc_feature

<222> (1)..(21)

<223> Primers

<400> 3

agaaaacaga tggcagaatg a

21

<210> 4

<211> 20

<212> DNA

<213> PRIMERS

<220>

<221> misc_feature

<222> (1)..(20)

<223> Primers

<400> 4

gcctgggcaa catagagaga

20

<210> 5

<211> 197

<212> DNA

<213> HUMAN

<400> 5

agaaaacaga tggcagaatg ataaactcaa tcatgttgat aaatatatta aatgtaaatg 60

gcttaaatat ccaactaaaa gactactaga atggattcta ttctattcta ttctattcra 120

ttctattcta ttctattcta ttctattcta ttctattcta ttctttttga gatgaagtct 180

ctctatgttg cccaggc 197

<210> 6
<211> 20
<212> DNA
<213> PRIMERS

<220>
<221> misc_feature
<222> (1)..(20)
<223> Primers

<400> 6
tccttcctca gtctttctgg

20

<210> 7
<211> 20
<212> DNA
<213> PRIMERS

<220>
<221> misc_feature
<222> (1)..(20)
<223> Primers

<400> 7
tgccatctgt tttctatttg

20

<210> 8
<211> 45

<212> DNA

<213> PROBES

<220>

<221> misc_feature

<222> (1)..(45)

<223> Probes

<400> 8

ctgttgtctt cagtgcagatt tcgcacagca tacaccaccc actgc

45

<210> 9

<211> 45

<212> DNA

<213> PROBES

<220>

<221> misc_feature

<222> (1)..(45)

<223> Probes

<400> 9

cactgcagag atgagaggtc cgtgagatgg aatctgaatg tgttc

45

<210> 10

<211> 27

<212> DNA

<213> PRIMERS

<220>

<221> misc_feature

<222> (1)..(27)

<223> Primers

<400> 10
gaagacaaat agaaaacaga tggcaga

27

<210> 11

<211> 52

<212> DNA

<213> PRIMERS

<220>

<221> misc_feature

<222> (1)..(52)

<223> Primers

<400> 11
tacgcatccc agtttgagac ggaatagaat agaatagaat agaatagaat ag

52

<210> 12

<211> 45

<212> DNA

<213> Mus musculus

<220>

<221> misc_feature

<222> (1)..(45)

<400> 12
ctgttgctctt cagtgagatt tcgcacagca tacaccaccc actgc

45

<210> 13

<211> 45

<212> DNA

<213> Mus musculus

<220>

<221> misc_feature

<222> (1)..(45)

<400> 13
cactgcagag atgagaggtc cgtgagatgg aatctgaatg tgttc

45

REFERENCES

[0113] The following references, to the extent that they provide exemplary procedural or other details supplementary to those set forth herein, are specifically incorporated herein by reference:

U. S. Patent 4,683,195

U. S. Patent 4,683,202

U. S. Patent 4,800,159

U. S. Patent 5,843,663

U. S. Patent 5,900,481

U. S. Patent 5,919,626

U. S. Patent 5,849,481

U. S. Patent 5,849,486

U. S. Patent 5,851,772

U. S. Patent 4,873,191

WO 90/07641

Affymax Technology

Bellus *et al.*, 1994

Bidichandani S.I., Ashizawa T. & Patel, P.I. The GAA triplet-repeat expansion in Friedreich ataxia interferes with transcription and may be associated with an unusual DNA structure. *Am. J. Hum. Genet.* 62, 111-121 (1998).

Breschel T.S. *et al.* A novel, heritable, expanding CTG repeat in an intron of the SEF2-1 gene on chromosome 18q21.1. *Hum. Mol. Genet.* 6, 1855-1863 (1997).

David G. *et al.* Cloning of the SCA7 gene reveals a highly unstable CAG repeat expansion. *Nature Genet.* 17, 65-70 (1997).

Dunham I. *et al.* The DNA sequence of human chromosome 22. *Nature* 402, 489-495 (1999).

Freifelder D, and Better M. (1982) Dialysis of small samples in agarose gels. *Anal*

Biochem. 123(1):83-5.

Higuchi R, Fockler C, Dollinger G, Watson R. (1993) Kinetic PCR analysis: real-time monitoring of DNA amplification reactions. *Biotechnology (N Y)*. Sep;11(9):1026-30.

Hogan, Beddington, Costantini and Long, eds. *Manipulating the Mouse Embryo; A Laboratory Manual* 2nd edition Cold Spring Harbor Laboratory Press, 1994.

Holmes S.E. *et al.* Expansion of a novel CAG trinucleotide repeat in the 5' region of PPP2R2B is associated with SCA12. *Nature Genet.* 23, 391-392 (1999).

Ikeuchi T., *et al.* A novel long and unstable CAG/CTG trinucleotide repeat on chromosome 17q. *Genomics.* 49, 321-326 (1998).

Imbert G. *et al.* Cloning of the gene for spinocerebellar ataxia 2 reveals a locus with high sensitivity to expanded CAG/glutamine repeats. *Nature Genet.* 14, 285-291 (1996).

Innis *et al.*, PCR Protocols, Academic Press, Inc., San Diego CA, 1990.

Kawaguchi Y. *et al.* CAG expansions in a novel gene for Machado-Joseph disease at chromosome 14q32.1. *Nature Genet.* 8, 221-228 (1994).

Kievits, *et al.*, *Cytogenet. Cell Genet.*, 53:134-136 (1990).

Koob M.D. *et al.* An untranslated CTG expansion causes a novel form of spinocerebellar ataxia (SCA8). *Nature Genet.* 21, 379-384 (1999).

Koob M.D. *et al.* Rapid cloning of expanded trinucleotide repeat sequences from genomic DNA. *Nature Genet.* 18, 72-75 (1998).

Matsuura T, Yamagata T, Burgess DL, *et al.* Large expansion of ATTCT pentanucleotide repeat in spinocerebellar ataxia type 10. *Nature Genet* 2000;26:191-194. Matsuura T. *et al.* Mapping of the gene for a novel spinocerebellar ataxia with pure cerebellar signs and epilepsy. *Ann. Neurol.* 45, 407-411 (1999).

Matsuura T., Watase K., Nagamitsu S., Zoghbi H.Y. & Ashizawa T. Fine mapping of the spinocerebellar ataxia type 10 region and search for a polyglutamine expansion. *Ann. Neurol.* 46, 480 (1999).

Nakamura K., *et al.* SCA17, a novel autosomal dominant cerebellar ataxia caused by an expanded polyglutamine in TATA-binding protein. *Hum. Mol. Genet.* 10:1441-1448 (2001).

O'Sullivan Smith *et al.*, Spinocerebellar Ataxia: Making an Informed Choice About Genetic Testing, National Institute on Disability and Rehabilitation Research, Washington, D.C. (1999).

Orr H.T. *et al.* Expansion of an unstable trinucleotide CAG repeat in spinocerebellar ataxia type 1. *Nature Genet.* 4, 221-226 (1993).

Palmiter RD, Brinster RL. (1985) Transgenic mice. *Cell.* 41(2):343-5.

Pulst S.M. *et al.* Moderate expansion of a normally biallelic trinucleotide repeat in spinocerebellar ataxia type 2. *Nature Genet.* 14, 269-276 (1996).

Sambrook *et al.*, Molecular Cloning, 2nd ed., Cold Spring Harbor Laboratory Press, CSH, 1.38-1.39, 1989.

Sanpei K. *et al.* Identification of the spinocerebellar ataxia type 2 gene using a direct identification of repeat expansion and cloning technique, DIRECT. *Nature Genet.* 14, 277-284 (1996).

Schalling M., Hudson T.J., Buetow K.H. & Housman D.E. Direct detection of novel expanded trinucleotide repeats in the human genome. *Nature Genet* 4, 135-139 (1993).

Schwartz, *et al.*, Cold Springs Harbor Symposium, Quantitative Biology, 47:189-195 (1982).

Trottier Y. *et al.* Polyglutamine expansion as a pathological epitope in Huntington's disease and four dominant cerebellar ataxias. *Nature* 378, 403-406 (1995).

Warner JP, Barron LH, Goudie D, et al. A general method for the detection of large CAG repeat expansions by fluorescent PCR. *J Med Genet* 1996;33:1022-1026. Wells R.D. & Warren S.T. Genetic instability and hereditary neurological diseases. San Diego, Academic Press (1998).

Wotta, *et al.*, *Am. J. of Human Genetics*, 46, 95-106 (1988)

Zhuchenko O. *et al.* Autosomal dominant cerebellar ataxia (SCA6) associated with small polyglutamine expansions in the alpha 1A-voltage-dependent calcium channel. *Nature Genet.* 15, 62-69 (1997).

Zu L., Figueroa K.P., Grewal R. & Pulst S.M. Mapping of a new autosomal dominant spinocerebellar ataxia to chromosome 22. *Am. J. Hum. Genet.* 64, 594-599 (1999).

Zuhlke C, *et al.* Different types of repeat expansion in the TATA-binding protein gene are associated with a new form of inherited ataxia. *Eur. J. Hum. Genet.* 9:160-164 (2001).

1000-924400